

On the Chernoff distance for asymptotic LOCC discrimination of bipartite quantum states

William Matthews*

Department of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.

Andreas Winter†

*Department of Mathematics, University of Bristol, Bristol BS8 1TW, U.K. and
Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, Singapore 117542*

(Dated: 22 October 2007)

Motivated by the recent discovery of a quantum Chernoff theorem for asymptotic state discrimination, we investigate the distinguishability of two bipartite mixed states under the constraint of local operations and classical communication (LOCC), in the limit of many copies. While for two pure states a result of Walgate *et al.* shows that LOCC is just as powerful as global measurements, data hiding states (DiVincenzo *et al.*) show that locality can impose severe restrictions on the distinguishability of even orthogonal states. Here we determine the optimal error probability and measurement to discriminate many copies of particular data hiding states (extremal $d \times d$ Werner states) by a linear programming approach. Surprisingly, the single-copy optimal measurement remains optimal for n copies, in the sense that the best strategy is measuring each copy separately, followed by a simple classical decision rule. We also put a lower bound on the bias with which states can be distinguished by separable operations.

I. INTRODUCTION

The non-classical nature of information represented in states of a bipartite quantum system is strikingly evident in the fact that, even allowing the experimenters (Alice and Bob) holding each of the subsystems to use local operations and *classical* communication (LOCC) freely, they cannot access the information as well as if they were in the same lab or could exchange quantum states. Thus, there is a specifically quantum obstruction to the distributed analysis of data and investigating this obstruction is a way of obtaining an understanding of the quantum nature of information.

The problem of LOCC discrimination of two or more states, has recently attracted quite considerable attention [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and what can be said at the very least that it is difficult. In the simplest example, the experimenters are given one of two states at random according to some probability distribution and their task is to unambiguously determine which state they have with the smallest possible error probability. Throughout this paper we'll use $P_{\text{err}}^X(\rho_1, \rho_2; p)$ to denote the minimum error with which the states ρ_1 and ρ_2 , with prior probabilities p and $1 - p$ respectively, can be distinguished by a POVM that can be implemented by operations in the class X . It will sometimes be convenient to refer to the optimal bias (over random guessing) instead of the optimal probability. This we define, as usual, by

$$B^X = 1 - 2P_{\text{err}}^X. \quad (1)$$

In this work we will talk about the well known classes of PPT-preserving (PPT) operations, separable (SEP) operations [14] and local operations with classical communication (LOCC), which obey the strict inclusions [15]

$$\text{LOCC} \subset \text{SEP} \subset \text{PPT} \subset \text{ALL}, \quad (2)$$

where ALL simply denotes the set of all possible global operations. Briefly, the POVMs which can be implemented by operations in these different classes can be characterized as follows: An LOCC POVM is one which can be implemented as a multi-round process where each round consists of a partial measurement of one party, which can depend on previously generated classical messages, and

*Electronic address: william.matthews@bris.ac.uk

†Electronic address: a.j.winter@bris.ac.uk

whose result is broadcast; A POVM is in SEP if and only if its elements can be written as positive linear combinations of product operators; A POVM can be implemented by PPT operations if and only if its constituent operators have positive partial transpose. The inclusion structure immediately implies the ordering

$$P_{\text{err}}^{\text{LOCC}} \geq P_{\text{err}}^{\text{SEP}} \geq P_{\text{err}}^{\text{PPT}} \geq P_{\text{err}}^{\text{ALL}} = \frac{1}{2} - \frac{1}{2} \|p\rho_1 - (1-p)\rho_2\|_1. \quad (3)$$

The final equality is the classic result of Helstrom and Holevo [16]. A similar closed form expression does not seem to exist for $P_{\text{err}}^{\text{LOCC}}$ or any of the other bipartite $P_{\text{err}}^{\text{X}}$.

Motivated by the recent development of a quantum Chernoff theorem [17], we are interested here in the asymptotic behaviour of the quantity $P_{\text{err}}^{\text{X}}(\rho_1^{\otimes n}, \rho_2^{\otimes n}; p)$ as the number of copies, n , goes to infinity. We can define the *Chernoff distance* with respect to a class of operations X , between the states ρ_1 and ρ_2 by

$$\xi^{\text{X}}(\rho_1, \rho_2) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log P_{\text{err}}^{\text{X}}(\rho_1^{\otimes n}, \rho_2^{\otimes n}; p). \quad (4)$$

(We note that the Chernoff distance is not strictly a distance since it does not obey the triangle inequality and that it is independent of the prior probabilities as long as they are both non-zero.)

In [17], it was determined that the (unconstrained) quantum Chernoff distance $\xi^{\text{ALL}}(\rho_1, \rho_2)$ is given by the formula (note the independence of p):

$$\xi^{\text{ALL}}(\rho_1, \rho_2) = -\min_{0 \leq s \leq 1} \log \text{Tr} \rho_1^{1-s} \rho_2^s. \quad (5)$$

This is a pleasantly straightforward generalisation of the classical Chernoff theorem for probability distributions, where for probability distribution vectors p and q

$$\xi(p, q) = -\min_{0 \leq s \leq 1} \log \sum_{i=1}^n p_i^{1-s} q_i^s. \quad (6)$$

It is useful to define yet another Chernoff distance on quantum states, for an even more restricted class of measurements than LOCC. Let $(M, \mathbb{1} - M)$ be the optimal single-copy LOCC POVM. $\xi^{\text{SC}}(\rho_1, \rho_2; p)$ is the classical Chernoff distance between the probability distributions on the outcome of this measurement when it is performed on ρ_1 or ρ_2 . (Outside the bipartite setting this notion was considered before by Kargin [18].) If we write

$$p_{j1} = \text{Tr}(M\rho_j), \quad p_{j2} = \text{Tr}((\mathbb{1} - M)\rho_j), \quad (7)$$

we can summarize the relationships between Chernoff distances we have defined as follows:

$$-\min_{0 \leq s \leq 1} \log \sum_{i=1}^2 p_{i1}^{1-s} p_{i2}^s = \xi^{\text{SC}} \leq \xi^{\text{LOCC}} \leq \xi^{\text{SEP}} \leq \xi^{\text{PPT}} \leq \xi^{\text{ALL}} = -\min_{0 \leq s \leq 1} \log \text{Tr} \rho_1^{1-s} \rho_2^s \quad (8)$$

Before proceeding with our main new results, we would like to make some general remarks about these quantities and describe some of the existing knowledge about them. One striking difference between global and local state discrimination can be seen in the effect of adding an ancilla. In the global case, this has no effect on our ability to distinguish between states, asymptotically or otherwise. That is, for any state τ

$$P_{\text{err}}^{\text{ALL}}(\rho_1, \rho_2; p) = P_{\text{err}}^{\text{ALL}}(\rho_1 \otimes \tau, \rho_2 \otimes \tau; p), \quad \xi^{\text{ALL}}(\rho_1, \rho_2; p) = \xi^{\text{ALL}}(\rho_1 \otimes \tau, \rho_2 \otimes \tau; p). \quad (9)$$

This is hardly surprising when one considers that the addition of any ancilla state is subsumed by the POVM formalism in the global case. In cases where our ability to distinguish between two states (of a $d \times d$ system, let's say) is worsened by restriction to LOCC, then we will indeed be helped by the provision of a $d \times d$ maximally entangled ancilla: by using it to teleport Alice's half to Bob (say), we have restored the ability to make global measurements and will be able to decrease the error probability accordingly. It is not always the case that the restriction to LOCC will impair our performance however. It was shown by Walgate *et al.* [1] (and generalized to non-orthogonal states

by Virmani *et al.* [2]) that LOCC can do just as well in distinguishing between two pure states as a global measurement can.

$$P_{\text{err}}^{\text{ALL}}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|; p) = P_{\text{err}}^{\text{LOCC}}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|; p). \quad (10)$$

Naturally, the corresponding Chernoff distances are also equal when both states are pure. Recently, Nathanson [19] has generalized this to the case of discriminating a mixed state from a pure state. He finds that under certain conditions on the fidelity of the states and the Schmidt coefficients of the pure state, $\xi^{\text{LOCC}}(\rho_1, \rho_2) = \xi^{\text{ALL}}(\rho_1, \rho_2)$, even though the single-copy error probabilities may differ.

From our perspective, it is more interesting to look at pairs of states where the LOCC constraint reduces our ability to distinguish them. In this paper we discuss an example of such a case. Let σ_d and α_d denote the completely symmetric and completely anti-symmetric Werner states in $d \times d$ dimensions, respectively (when d is a power of two, these are the states used by DiVincenzo *et al.* [20] for “data hiding”; see also [21]). In this paper we calculate the Chernoff distance between these states, $\xi^{\text{LOCC}}(\sigma_d, \alpha_d)$, and to do so, we actually give an expression for $P_{\text{err}}^{\text{LOCC}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p)$.

The rest of this paper is organized as follows: In the next section we present an LOCC protocol which puts an upper bound on $P_{\text{err}}^{\text{LOCC}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p)$. In section III, we formulate the minimization of the error which can be achieved by PPT operations as a linear program, and by solving the dual program show that the LOCC upper bound is also a lower bound on $P_{\text{err}}^{\text{PPT}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p)$ and hence on $P_{\text{err}}^{\text{LOCC}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p)$, thus proving the optimality of our LOCC protocol, and allowing us to calculate the Chernoff distance. In section IV, we prove a lower bound on $B^{\text{SEP}}(\rho_1, \rho_2; p)$ in terms of $B^{\text{ALL}}(\rho_1, \rho_2; p)$, after which we conclude.

To describe asymptotic behaviours we will use ‘Big-O’ notation (including Θ, Ω and \sim). If X is an operator on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, we use X^T to denote its partial transpose, which is defined (for some orthonormal product basis $\{|i\rangle_A \otimes |j\rangle_B\}$) by

$$|i\rangle_A \otimes |j\rangle_B \langle k|_A \otimes \langle l|_B^T = |i\rangle_A \otimes |l\rangle_B \langle k|_A \otimes \langle j|_B. \quad (11)$$

II. LOCC DISCRIMINATION PROTOCOL

Proposition 1. *There is an LOCC protocol (requiring only one-way communication) which demonstrates that $P_{\text{err}}^{\text{LOCC}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) \leq \min\left(p\left(\frac{d-1}{d+1}\right)^n, 1-p\right)$.*

Proof. Alice and Bob take each copy in turn and measure in the computational basis. They share their results. If they recorded different results for every copy then they guess that they have the anti-symmetric state. Otherwise, they have obtained the same result for at least one state and they know with certainty that they share the symmetric state.

For a single copy, the POVM implemented by this measurement is

$$\left\{ G_d = \sum_{i \neq j}^{d-1} |ij\rangle\langle ij|, \mathbb{1} - G_d = \sum_{i=0}^{d-1} |ii\rangle\langle ii| \right\}. \quad (12)$$

Because the states to be distinguished are both $U \otimes U$ -invariant, it is convenient to apply the twirl operation to the two operators in the POVM and it also emphasizes the symmetry of the states that are to be distinguished. After doing so we have the following single-copy POVM of equal performance:

$$\left\{ M_d = \frac{d-1}{d+1} \Pi_s + \Pi_a, \mathbb{1} - M_d = \frac{2}{d+1} \Pi_s \right\}, \quad (13)$$

where Π_s and Π_a are the projections onto the symmetric and anti-symmetric subspaces, respectively. The POVM element M_d corresponds to Alice and Bob having different measurement outcomes on a single copy. For n copies the POVM is

$$\{M_d^{\otimes n}, \mathbb{1} - M_d^{\otimes n}\}, \quad (14)$$

since $M_d^{\otimes n}$ corresponds to Alice and Bob getting different outcomes for every copy they measure. Let A_k denote the sum of all elements of $\{\Pi_s, \Pi_a\}^{\otimes n}$ which have k copies of Π_a . Expanding in terms of the $n + 1$ orthogonal projection operators $\{A_0, \dots, A_n\}$, we find that

$$M_d^{\otimes n} = \sum_{k=0}^n \left(\frac{d-1}{d+1} \right)^{n-k} A_k. \quad (15)$$

$$P_{\text{err}} = p \text{Tr} (M_d^{\otimes n} \sigma_d^{\otimes n}) + (1-p) \text{Tr} ((\mathbb{1} - M_d^{\otimes n}) \alpha_d^{\otimes n}), \quad (16)$$

where the first term is the probability that Alice and Bob have the symmetric state and mistake it for the anti-symmetric state and the second term is the probability that they share the anti-symmetric state and mistake it for the symmetric state.

Substituting (15) into (16) and using the fact that $\sigma_d^{\otimes n} \propto A_0$ and $\alpha_d^{\otimes n} \propto A_n$, we obtain

$$P_{\text{err}} = p \text{Tr} \left(\left(\frac{d-1}{d+1} \right)^n A_0 \sigma_d^{\otimes n} \right) + (1-p) \text{Tr} ((1 - A_n) \alpha_d^{\otimes n}) = p \left(\frac{d-1}{d+1} \right)^n. \quad (17)$$

If $P_{\text{err}} > 1 - p$ then we will do better to simply guess that we have the symmetric state all the time. Adding this proviso to our strategy, we obtain the desired result. \square

Remark 2. We note that the second term in the expression for the error probability is zero, meaning that all the error is due to the case where the symmetric state is mistaken for the anti-symmetric state. This is just what we would expect given that our protocol reports that we have a symmetric state only when it is certain that we have one.

We shall now show that (17) is the optimum error probability that can be achieved using LOCC by showing that it is the best that can be achieved even if we use the larger class of measurements that can be implemented using PPT preserving operations.

III. OPTIMAL PPT PRESERVING POVM

We shall first formulate the minimisation of the error probability over PPT preserving POVMs [14] as linear programming problem (see [22], for instance) by taking advantage of the symmetries of the states we wish to distinguish. We will then show that there is a solution to the dual linear program which lower bounds the error probability to exactly that achieved by the LOCC procedure given above.

The states $\alpha_d^{\otimes n}$ and $\sigma_d^{\otimes n}$ are invariant under permutations of the copies and under biunitary transformations of the individual copies. We can assume therefore that our two POVM elements have the same symmetries (this is a trick that was used before in [24] to solve a relative entropy minimisation problem). The most general operator with these symmetries is a linear combination of the $n + 1$ operators A_k which we defined above, so we write our POVM as:

$$\left\{ \sum_{k=0}^n x_k A_k, \sum_{k=0}^n (1 - x_k) A_k \right\}. \quad (18)$$

The probability of error is given by

$$P_{\text{err}} = p \text{Tr} \left(\sum_{k=0}^n x_k A_k \sigma_d^{\otimes n} \right) + (1-p) \text{Tr} \left(\sum_{k=0}^n (1 - x_k) A_k \alpha_d^{\otimes n} \right) = (1-p) + p \left(x_0 - \frac{1-p}{p} x_n \right). \quad (19)$$

The constraints

$$x_k \geq 0 \quad \text{for } k = 0, \dots, n, \quad (20)$$

$$x_k \leq 1 \quad \text{for } k = 0, \dots, n \quad (21)$$

are necessary and sufficient to ensure that the two operators do in fact comprise a POVM.

The partial transpose of the flip operator F is equal to $d\Phi_d$, where $\Phi_d = \frac{1}{d} \sum_{i,j=0}^{d-1} |ii\rangle\langle jj|$ is the maximally entangled state. Since $\Pi_s = (\mathbb{1} + F)/2$ and $\Pi_a = (\mathbb{1} - F)/2$, we have

$$\Pi_s^\Gamma = \frac{1}{2}(\mathbb{1} + d\Phi_d) = \frac{1}{2}((\mathbb{1} - \Phi_d) + (1 + d)\Phi_d), \quad (22)$$

$$\Pi_a^\Gamma = \frac{1}{2}(\mathbb{1} - d\Phi_d) = \frac{1}{2}((\mathbb{1} - \Phi_d) + (1 - d)\Phi_d), \quad (23)$$

so the operators A_k^Γ can be written as linear combinations of operators from the set of 2^n orthogonal operators $\{(\mathbb{1} - \Phi_d), \Phi_d\}^{\otimes n}$.

Let S_k^n denote the subset of strings in $\{0, 1\}^N$ which have exactly k ones. Then,

$$\begin{aligned} A_k^\Gamma &= 2^{-n} \sum_{v \in S_k^n} \bigotimes_{i=1}^n ((\mathbb{1} - \Phi_d) + (1 + (-1)^{v_i} d) \Phi_d) \\ &= 2^{-n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} (1+d)^j (1-d)^{l-j} T_l, \end{aligned} \quad (24)$$

where T_l is the sum over all elements of $\{(\mathbb{1} - \Phi_d), \Phi_d\}^{\otimes n}$ which have l copies of Φ_d .

A POVM is PPT preserving if and only if all of the operators that comprise it have positive partial transpose [14]. A necessary and sufficient condition for the POVM to be PPT preserving is therefore given by the following inequalities

$$\sum_{k=0}^n x_k \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} (1-d)^j (1+d)^{l-j} \geq 0 \quad \text{for } l = 0, \dots, n, \quad (25)$$

$$\sum_{k=0}^n (1 - x_k) \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} (1-d)^j (1+d)^{l-j} \geq 0 \quad \text{for } l = 0, \dots, n. \quad (26)$$

Let Q be an $(n+1) \times (n+1)$ matrix with elements

$$Q_{lk} = \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} (1-d)^j (1+d)^{l-j}. \quad (27)$$

We note that

$$\begin{aligned} \sum_{k=0}^n Q_{lk} &= (1+d)^l \sum_{m=0}^{n-l} \binom{n-l}{m} \sum_{j=0}^l \binom{l}{j} \left(\frac{1-d}{1+d}\right)^j \\ &= (1+d)^l \left(1 + \frac{1-d}{1+d}\right)^l \sum_{m=0}^{n-l} \binom{n-l}{m} \\ &= (1+d)^l \left(\frac{2}{1+d}\right)^l 2^{n-l} = 2^n. \end{aligned} \quad (28)$$

Defining the vectors c and b as follows

$$c_i = \delta_{0i} - \frac{1-p}{p} \delta_{ni}, \quad (29)$$

$$b_i = \begin{cases} 0 & \text{for } i = 0, \dots, n, \\ -2^n & \text{for } i = n+1, \dots, 2n+1, \\ -1 & \text{for } i = 2n+2, \dots, 3n+2, \end{cases} \quad (30)$$

we can write the optimisation in standard linear programming form

$$\min_x \{c^T \cdot x \mid P \cdot x \geq b, x \geq 0\} \text{ where } P = \begin{pmatrix} Q \\ -Q \\ -\mathbb{1} \end{pmatrix}. \quad (31)$$

Writing (19) in terms of the objective function $c^T \cdot x$, we see that the POVM corresponding the vector x has error probability

$$P_{\text{err}}(x) = (1 - p) + pc^T \cdot x. \quad (32)$$

Proposition 3. *The probability of error for a PPT preserving POVM to distinguish $\sigma_d^{\otimes n}$ and $\alpha_d^{\otimes n}$ with prior probabilities p and $1 - p$, $P_{\text{err}}^{\text{PPT}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p)$, is bounded below by $\min\left(p\left(\frac{d-1}{d+1}\right)^n, 1 - p\right)$.*

Proof. The linear program dual to (31) is just

$$\max_y \{b^T \cdot y \mid P^T \cdot y \leq c, y \geq 0\}. \quad (33)$$

Indeed, the duality of linear programs tells that for any primal feasible point x and any dual feasible point y

$$c^T \cdot x \geq b^T \cdot y, \quad (34)$$

so any dual feasible point y gives us a lower bound on the error probability:

$$P_{\text{err}}^{\text{PPT}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) \geq (1 - p) + pb^T \cdot y. \quad (35)$$

It is convenient to write y as the direct sum of three $(n + 1)$ -dimensional vectors $y = u \oplus v \oplus w$ so that we can rewrite the dual program as

$$\max_y \left\{ -2^n \sum_{i=0}^n v_i - \sum_{i=0}^n w_i \mid u \geq 0, v \geq 0, w \geq 0, Q^T \cdot u - Q^T \cdot v - w \leq c \right\}. \quad (36)$$

Consider the point $y^* = u^* \oplus v^* \oplus w^*$ defined by

$$u_i^* = \binom{n}{i} \frac{(d-1)^{n-i}((d+1)^i - (1-d)^i)}{(2d)^n(d+1)^i}, \quad (37)$$

$$v_i^* = 0, \quad (38)$$

$$w_i^* = \begin{cases} 0 & \text{for } i = 0, \dots, n-1, \\ \max\left(\frac{1-p}{p} - \left(\frac{d-1}{d+1}\right)^n, 0\right) & \text{for } i = n. \end{cases} \quad (39)$$

We show that the point y^* is dual feasible in Appendix A. The dual objective function at this point is

$$-2^n \sum_{i=0}^n v_i^* - \sum_{i=0}^n w_i^* = -w_n^* = \min\left(\left(\frac{d-1}{d+1}\right)^n - \frac{1-p}{p}, 0\right), \quad (40)$$

so, substituting y^* into (35), we obtain the bound:

$$P_{\text{err}}^{\text{PPT}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) \geq \min\left(p\left(\frac{d-1}{d+1}\right)^n, 1 - p\right). \quad (41)$$

□

Corollary 4. *Substituting the results of Proposition 1 and Proposition 3 into (3), we have shown that*

$$P_{\text{err}}^{\text{PPT}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) = P_{\text{err}}^{\text{SEP}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) = P_{\text{err}}^{\text{LOCC}}(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}; p) = \min\left(p\left(\frac{d-1}{d+1}\right)^n, 1 - p\right). \quad (42)$$

Substituting into the definition of the Chernoff information for each class of operations and noting that each copy is measured separately in the optimal strategy, we obtain our main result:

Theorem 5. *Whenever $0 < p < 1$, we have*

$$\xi^{\text{PPT}}(\sigma_d, \alpha_d) = \xi^{\text{SEP}}(\sigma_d, \alpha_d) = \xi^{\text{LOCC}}(\sigma_d, \alpha_d) = \xi^{\text{SC}}(\sigma_d, \alpha_d) = \log \frac{d+1}{d-1} \sim \frac{2 \log e}{d-1}. \quad (43)$$

IV. A LOWER BOUND ON BIAS FOR SINGLE-COPY SEPARABLE MEASUREMENTS

The fact that $\xi^{\text{LOCC}}(\sigma_d, \alpha_d) = \xi^{\text{SC}}(\sigma_d, \alpha_d)$ shows that our ability to distinguish the extremal Werner states cannot be improved by measurements which are entangled across copies. This is the least favorable many-copy behaviour possible. It would be interesting to know if the single-copy error probability for these states also has the worst kind of scaling with dimension. In terms of bias, we have shown that

$$\frac{B^{\text{LOCC}}(\sigma_d, \alpha_d; p)}{B^{\text{ALL}}(\sigma_d, \alpha_d; p)} = \Theta\left(\frac{1}{d}\right). \quad (44)$$

Is $1/d$ an asymptotic lower bound whatever states we choose? If we relax the LOCC constraint and allow separable operations then we can show that it is.

Proposition 6. *If ρ_1 and ρ_2 are bipartite states on a system of overall dimension D , then*

$$B^{\text{SEP}}(\rho_1, \rho_2; p) \geq \frac{1}{2\sqrt{D}} B^{\text{ALL}}(\rho_1, \rho_2; p). \quad (45)$$

Proof. We know that the optimal error probability for global measurements is given by the Holevo-Helstrom POVM, the elements of which are generally not even PPT. It was shown by Barnum and Gurvits [23] that every Hermitian operator in the ball centred on the identity, with radius one in the Hilbert-Schmidt norm is separable. If we add to each element of the Holevo-Helstrom POVM the minimum amount of the identity operator necessary to put the resulting operator inside this ball, and normalize the POVM, we obtain the *separable* POVM

$$\left\{ \frac{1}{2} \left(\mathbb{1} + \frac{M}{\|M\|_2} \right), \frac{1}{2} \left(\mathbb{1} - \frac{M}{\|M\|_2} \right) \right\} \quad (46)$$

where M is the projector onto the support of the positive part of $(1-p)\rho_2 - p\rho_1$ if $p \leq 1/2$ (and minus one times the projector onto the support of the negative part otherwise). This POVM yields the error probability

$$P_{\text{err}} = \frac{1}{2} \left(1 - \frac{1}{2\|M\|_2} (|1-2p| + \|(1-p)\rho_2 - p\rho_1\|_1) \right). \quad (47)$$

Using the fact that $\|M\|_2 \leq \sqrt{D}$, we get the bound

$$\|(1-p)\rho_2 - p\rho_1\|_1 = B^{\text{ALL}} \geq B^{\text{PPT}} \geq B^{\text{SEP}} \geq \frac{1}{2\sqrt{D}} \|(1-p)\rho_2 - p\rho_1\|_1 = \frac{1}{2\sqrt{D}} B^{\text{ALL}}. \quad (48)$$

□

So, for states of a $d \times d$ system: $B^{\text{SEP}}/B^{\text{ALL}} \in \Omega(1/d)$. This result, combined with our result for the data hiding states, leads us to conjecture that

Conjecture 7. *For states on a $d \times d$ system,*

$$\frac{B^{\text{LOCC}}}{B^{\text{ALL}}} \geq \Omega\left(\frac{1}{d}\right). \quad (49)$$

To put the insights and conjecture above into a different and wider perspective, let us look at the biases B^X for the particular value $p = \frac{1}{2}$:

$$B^X(\rho_1, \rho_2) := B^X\left(\rho_1, \rho_2; \frac{1}{2}\right), \quad (50)$$

for which, by definition, it is clear that it is symmetric: $B^X(\rho_1, \rho_2) = B^X(\rho_2, \rho_1)$. Furthermore, for all the classes X considered in the introduction, $B^X(\rho_1, \rho_2) = 0$ if and only if $\rho_1 = \rho_2$. Indeed, the

B^X are all metrics, as they obey the triangle inequality: $B^X(\rho_1, \rho_3) \leq B^X(\rho_1, \rho_2) + B^X(\rho_2, \rho_3)$ for any states ρ_1, ρ_2 and ρ_3 . To be more precise, they derive from operator norms $\|\cdot\|_X$, defined on trace-free hermitian operators:

$$B^X(\rho_1, \rho_2) = \left\| \frac{1}{2}(\rho_1 - \rho_2) \right\|_X, \text{ with } \|M\|_X = \sup_{\text{POVM } (M_i)_{i \in X}} \sum_i |\text{Tr } M M_i|, \quad (51)$$

We note that the supremum in (51) is always attained by a POVM with two elements (one with $\text{Tr}(M M_1) \geq 0$ and the other with $\text{Tr}(M M_2) = -\text{Tr}(M M_1) \leq 0$).

For example by Helstrom's theorem [16], $B^X(\rho_1, \rho_2) = \left\| \frac{1}{2}(\rho_1 - \rho_2) \right\|_1$, so $\|\cdot\|_{\text{ALL}} = \|\cdot\|_1$.

Of course, all norms on finite-dimensional spaces are equivalent up to constant factors. Eq. (48) translates into the ordering of norms

$$\|M\|_1 = \|M\|_{\text{ALL}} \geq \|M\|_{\text{PPT}} \geq \|M\|_{\text{SEP}} \geq \sqrt{\frac{1}{D}} \|M\|_{\text{ALL}}, \quad (52)$$

and Conjecture 7 can be expressed as $\|M\|_{\text{LOCC}} \geq \Omega\left(\frac{1}{d}\right) \|M\|_{\text{ALL}}$ for $d \times d$ systems. Note that the existence of data hiding states implies that this would be essentially best possible, as for $M = \frac{1}{2}(\alpha_d - \sigma_d)$,

$$\|M\|_{\text{LOCC}} \leq \|M\|_{\text{SEP}} \leq \|M\|_{\text{PPT}} = \frac{2}{d+1} \|M\|_{\text{ALL}}. \quad (53)$$

V. DISCUSSION

We have calculated the Chernoff distance between the extremal $d \times d$ Werner states, under the constraint of LOCC operations, for all values of d . This is the first time the LOCC Chernoff distance has been calculated for states where it differs from the unconstrained Chernoff distance. In this case, we have also been able to calculate the smallest error probability that can be achieved by LOCC for any finite number of copies. The solution has at least two remarkable features: First, the error probability is – up to constant factors – equal to the n -th power of the single-copy error probability, showing that in a sense n copies don't give disproportionate advantage over one copy, in this case. Secondly, even the optimal n -copy measurement reflects this structurally; namely, it can be implemented by measuring the single-copy optimal POVM n times, followed by a trivial classical post-processing. As discussed in the introduction, this is a “worst-case” strategy for many copies. Both of these properties distinguish the solution from what is to be expected in the quantum Chernoff problem: e.g., discriminating two (non-orthogonal) pure states has a very simple optimal strategy, but for n copies (which is also a problem of discriminating two pure states) this strategy is highly collective over the n systems. Also, in general, even classically, the error probability shows only an asymptotically exponential decay, but here it is exactly exponential.

Our result also leads to a number of further questions. An extension of the work which we are currently considering is to see if we can find Chernoff bounds for the discrimination of pairs of general Werner states. Preliminary and ongoing investigations suggest that some interesting effects occur when at least one state is non-extremal. Also, as discussed above, it would be interesting to know how close to “worst possible” is our example in terms of comparing LOCC to unrestricted measurements? That is, we would like to resolve our Conjecture 7 on the single-copy LOCC bias.

Acknowledgments

WM acknowledges support from the U.K. EPSRC; AW was supported through an Advanced Research Fellowship of the U.K. EPSRC, the EPSRC's “QIP IRC”, and the European Commission IP “QAP”. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

The authors would like to acknowledge useful discussions with Keiji Matsumoto, Chris King and Michael Nathanson and to thank Aram Harrow for a stimulating conversation on the design of the

optimally discriminating POVM.

-
- [1] J. Walgate, A. J. Short, L. Hardy and V. Vedral, “Local Distinguishability of Multipartite Orthogonal Quantum States”, *Phys. Rev. Lett.* **8**(23):4972-4975 (2000); arXiv:quant-ph/0007098.
 - [2] S. Virmani, M. F. Sacchi, M. B. Plenio and D. Markham, “Optimal local discrimination of two multipartite pure states”, *Phys. Lett. A* **288**:62-68 (2001); arXiv:quant-ph/0102073.
 - [3] H. Fan, “Distinguishing bipartite states by local operations and classical communication”, *Phys. Rev. A* **75**:014305 (2007).
 - [4] M. Hayashi, D. Markham, M. Murao, M. Owari and S. Virmani, “Bounds on Multipartite Entangled Orthogonal State Discrimination Using Local Operations and Classical Communication”, *Phys. Rev. Lett.* **96**:040501 (2006).
 - [5] J. Watrous, “Bipartite Subspaces Having No Bases Distinguishable by Local Operations and Classical Communication”, *Phys. Rev. Lett.* **95**:080505 (2005).
 - [6] M. Horodecki, J. Oppenheim, A. Sen(De) and U. Sen, “Distillation Protocols: Output Entanglement and Local Mutual Information”, *Phys. Rev. Lett.* **93**:170503 (2004).
 - [7] S. Ghosh, P. Joag, G. Kar, S. Kunkri, and A. Roy, “Locally accessible information and distillation of entanglement”, *Phys. Rev. A* **71**:012321 (2005).
 - [8] J. Walgate and L. Hardy, “Nonlocality, Asymmetry, and Distinguishing Bipartite States”, *Phys. Rev. Lett.* **89**:147901 (2002).
 - [9] B. Groisman and B. Reznik, “Measurements of semilocal and nonmaximally entangled states”, *Phys. Rev. A* **66**(2):022110 (2002).
 - [10] A. Chefles, “Condition for unambiguous state discrimination using local operations and classical communication”, *Phys. Rev. A* **69**:050307(R) (2004).
 - [11] M. Hayashi, K. Matsumoto and Y. Tsuda, “A study of LOCC-detection of a maximally entangled state using hypothesis testing”, *J. Phys. A: Math. Gen.* **39**:14427-14446 (2006); arXiv:quant-ph/0504203.
 - [12] C. King and D. Matysiak, “Reliably distinguishing states in qutrit channels using one-way LOCC”, arXiv:quant-ph/0510004 (2005).
 - [13] M. Nathanson, “Distinguishing bipartite orthogonal states using LOCC: Best and worst cases”, *J. Math. Phys.* **46**:062103 (2005); arXiv:quant-ph/0411110.
 - [14] E. M. Rains, “A semidefinite program for distillable entanglement”, *IEEE Trans. Inf. Theory*, **47**(7):2921-2933 (2001).
 - [15] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin and W. K. Wootters, “Quantum nonlocality without entanglement”, *Phys. Rev. A* **59**(2):1070-1091 (1999).
 - [16] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, New York, (1976).
 - [17] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, Ll. Masanes, A. Acín and F. Verstraete, “Discriminating States: The Quantum Chernoff Bound”, *Phys. Rev. Lett.* **98**:160501 (2007); arXiv:quant-ph/0610027. M. Nussbaum and A. Szkola, “A lower bound of Chernoff type for symmetric quantum hypothesis testing”, arXiv:quant-ph/0607216 (2006).
 - [18] V. Kargin, “On the Chernoff Bound for Efficiency of Quantum Hypothesis Testing”, *Ann. Stat.* **33**(2): 959-976 (2005).
 - [19] M. Nathanson, “Distinguishing a pure state from an arbitrary mixed state using LOCC”, in preparation (2007).
 - [20] D. P. DiVincenzo, D. W. Leung and B. M. Terhal, “Quantum data hiding”, *IEEE Trans. Inf. Theory* **48**(3):580-598 (2002); arXiv:quant-ph/0103098.
 - [21] T. Eggeling and R. F. Werner, “Hiding Classical Data in Multipartite Quantum States”, *Phys. Rev. Lett.* **89**:097905 (2002).
 - [22] A. Schrijver, *Theory of Linear and Integer Programming*, John Wiley and Sons (1998).
 - [23] H. Barnum, L. Gurvits, “Largest separable balls around the maximally mixed bipartite quantum state”, *Phys. Rev. A* **66**, 062311 (2002).
 - [24] K. M. R. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. S. Virmani and B. De Moor, *Phys. Rev. Lett.* **87**:217902 (2002); arXiv:quant-ph/0205025.

APPENDIX A: PROOF OF DUAL FEASIBILITY

We note that $u_i^* \geq 0$ for $i = 0, \dots, n$:

$$\begin{aligned} u_k^* &= (2d)^{-n} \binom{n}{k} \left((d-1)^{n-k} - (-1)^k \frac{(d-1)^n}{(d+1)^k} \right) \\ &\geq (2d)^{-n} \binom{n}{k} \left((d-1)^{n-k} - \frac{(d-1)^n}{(d+1)^k} \right) \\ &= \binom{n}{k} \left(\frac{d-1}{2d} \right)^n \left(\frac{1}{(d-1)^k} - \frac{1}{(d+1)^k} \right) \geq 0. \end{aligned} \quad (\text{A1})$$

It is obvious that $v^* \geq 0$ and $w^* \geq 0$, so the first three inequalities of (36) are satisfied.

We now show that the remaining inequality,

$$Q^T \cdot u - Q^T \cdot v - w \leq c, \quad (\text{A2})$$

is also satisfied:

$$\begin{aligned} (Q^T \cdot u^*)_k &= \frac{(d-1)^n}{(2d)^n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} \binom{n}{l} (1-d)^j (1+d)^{l-j} \frac{(d+1)^l - (1-d)^l}{(d-1)^l (d+1)^l} \\ &= s_1(d, n; l) - s_2(d, n; l), \end{aligned} \quad (\text{A3})$$

where

$$\begin{aligned} s_1(d, n; k) &= \frac{(d-1)^n}{(2d)^n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} \binom{n}{l} (1-d)^j (1+d)^{l-j} \frac{(d+1)^l}{(d-1)^l (d+1)^l} \\ &= \frac{(d-1)^n}{(2d)^n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} \binom{n}{l} (-1)^j \left(\frac{d+1}{d-1} \right)^{l-j}, \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} s_2(d, n; k) &= \frac{(d-1)^n}{(2d)^n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} \binom{n}{l} (1-d)^j (1+d)^{l-j} \frac{(1-d)^l}{(d-1)^l (d+1)^l} \\ &= \frac{(d-1)^n}{(2d)^n} \sum_{l=0}^n \sum_{0 \leq j \leq l, k} \binom{n-l}{k-j} \binom{l}{j} \binom{n}{l} (-1)^{j+l} \left(\frac{d-1}{d+1} \right)^j. \end{aligned} \quad (\text{A5})$$

Defining $m = l - j$ we can rewrite the first sum (A4) as

$$\begin{aligned} s_1(d, n; k) &= \frac{(d-1)^n}{(2d)^n} \sum_{m=0}^{n-k} \sum_{j=0}^k \binom{n-(m+j)}{k-j} \binom{m+j}{j} \binom{n}{m+j} (-1)^j \left(\frac{d+1}{d-1} \right)^m \\ &= \frac{(d-1)^n}{(2d)^n} \sum_{m=0}^{n-k} \sum_{j=0}^k \frac{n!}{(k-j)!(n-(m+k))!m!j!} (-1)^j \left(\frac{d+1}{d-1} \right)^m \\ &= \frac{(d-1)^n}{(2d)^n} \sum_{m=0}^{n-k} \sum_{j=0}^k \frac{n!}{(n-m)!m!} \frac{(n-m)!}{((n-m)-k)!k!} \frac{k!}{(k-j)!j!} (-1)^j \left(\frac{d+1}{d-1} \right)^m \\ &= \frac{(d-1)^n}{(2d)^n} \sum_{m=0}^{n-k} \binom{n}{m} \binom{n-m}{k} \left(\frac{d+1}{d-1} \right)^m \sum_{j=0}^k \binom{k}{j} (-1)^j. \end{aligned} \quad (\text{A6})$$

The sum over j is 0 except when $k = 0$, so

$$s_1(d, n; k) = \delta_{0k} \frac{(d-1)^n}{(2d)^n} \sum_{m=0}^n \binom{n}{m} \left(\frac{d+1}{d-1} \right)^m = \delta_{0k} \frac{(d-1)^n}{(2d)^n} \left(1 + \frac{d+1}{d-1} \right)^n = \delta_{0k}. \quad (\text{A7})$$

Making the same change of variables ($m = l - j$) in (A5), we obtain

$$\begin{aligned}
s_2(d, n; k) &= \frac{(d-1)^n}{(2d)^n} \sum_{j=0}^k \sum_{l=j}^{n+j-k} \binom{n}{l} \binom{n-l}{k-j} \binom{l}{j} (-1)^{j+l} \left(\frac{d-1}{d+1} \right)^j \\
&= \frac{(d-1)^n}{(2d)^n} \sum_{j=0}^k \sum_{m=0}^{n-k} \binom{n}{m+j} \binom{n-(m+j)}{k-j} \binom{m+j}{j} (-1)^j (-1)^{m+j} \left(\frac{d-1}{d+1} \right)^j \\
&= \frac{(d-1)^n}{(2d)^n} \sum_{j=0}^k \sum_{m=0}^{n-k} \frac{n!}{(k-j)!(n-(m+k))!m!j!} (-1)^{2j} (-1)^m \left(\frac{d-1}{d+1} \right)^j \\
&= \frac{(d-1)^n}{(2d)^n} \sum_{j=0}^k \sum_{m=0}^{n-k} \frac{n!}{(n-k)!k!} \frac{(n-k)!}{((n-k)-m)!m!} \frac{k!}{(k-j)!j!} (-1)^m \left(\frac{d-1}{d+1} \right)^j \\
&= \frac{(d-1)^n}{(2d)^n} \binom{n}{k} \sum_{m=0}^{n-k} \binom{n-k}{m} (-1)^m \sum_{j=0}^k \binom{k}{j} \left(\frac{d-1}{d+1} \right)^j \\
&= \delta_{nk} \frac{(d-1)^n}{(2d)^n} \binom{n}{k} \left(1 + \frac{d-1}{d+1} \right)^n = \delta_{nk} \left(\frac{d-1}{d+1} \right)^n.
\end{aligned} \tag{A8}$$

Substituting (A7) and (A8) into (A3) we find that $(Q^T \cdot u^*)_k = \delta_{0k} - \delta_{nk} \left(\frac{d-1}{d+1} \right)^n$, so the constraint (A2) is satisfied:

$$(Q^T \cdot u^* - Q^T \cdot v^* - w^*)_k = \delta_{0k} - \delta_{nk} \left(\frac{d-1}{d+1} \right)^n - \max \left(r - \left(\frac{d-1}{d+1} \right)^n, 0 \right) \delta_{nk} \leq c_k. \tag{A9}$$